



Chung How, JTH., & Bull, D. R. (1999). *Robust image and video coding with pyramid vector quantisation*. 332 - 335.
<https://doi.org/10.1109/ISCAS.1999.780009>

Peer reviewed version

Link to published version (if available):
[10.1109/ISCAS.1999.780009](https://doi.org/10.1109/ISCAS.1999.780009)

[Link to publication record in Explore Bristol Research](#)
PDF-document

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

ROBUST IMAGE AND VIDEO CODING WITH PYRAMID VECTOR QUANTISATION

J.T.H. Chung-How and D.R. Bull

Image Communications Group, Centre for Communications Research
University of Bristol, Woodland Road, Bristol BS8 1UB, U.K.

ABSTRACT

Most current image and video coding standards use variable length codes to achieve compression, which renders the compressed bitstream very sensitive to channel errors. In this paper, image and video coders based on Pyramid Vector Quantisation (PVQ) and using only fixed length codes are proposed. Still image coders using PVQ in conjunction with DCT and wavelet techniques are described and their robustness to random channel errors are investigated. This work is then extended to video coding. A novel fixed rate motion compensated wavelet/PVQ video coder suitable for low bit-rate applications and generating only fixed length codewords is presented. Its compression performance and robustness to random bit errors is compared with H.263.

1. INTRODUCTION

The popularity and widespread availability of mobile wireless networks, combined with the emergence of powerful portable digital computers, have resulted in an increased interest in image and video transmission over wireless networks. Potential applications include mobile videoconferencing, remote surveillance and monitoring, and emergency police and medical services. Due to the limited bandwidth of wireless networks, efficient coding algorithms are required to compress the image and video for transmission. However, wireless channels typically suffer from multipath and fading, resulting in unpredictable bit-errors in the received signal. Therefore, in order to transmit images or video over a wireless network, the coding algorithm must deliver good compression and be robust to errors.

Existing image and video coding standards such as JPEG, H.263 and MPEG use variable length coding in the form of Huffman or arithmetic coding in order to achieve compression. The problem with variable length codes is that a single bit-error in a codeword is likely to cause a wrong codeword of the wrong length to be decoded, resulting in loss of codeword synchronisation. Thus a single bit-error can propagate from one codeword to the next and result in catastrophic failure at the decoder. Error propagation can be limited by using synchronisation codewords, which are special variable length codes which enables the decoder to regain codeword synchronisation provided the synchronisation codeword is received error-free. Huffman codes can be designed to include at least one synchronisation codeword without any loss in compression efficiency [1], but this can still result in a correct codeword being decoded in the wrong location. Alternatively, explicit synchronisation codewords can be inserted at fixed locations in the bitstream. However, such synchronisation codewords result in considerable overhead and cannot be used too frequently. Reversible or two-way decodable variable length codes have also been investigated [2]. In [3], a technique known

as error resilient entropy coding (EREC) is proposed which rearranges variable length blocks into fixed length slots with minimal redundancy. The EREC can be applied to still image and video coding and provides considerable robustness to both burst and random errors.

A different approach to limiting error propagation is to avoid using variable length codes and use fixed length codes instead. However, fixed length codes are optimum only if all codewords are equiprobable. Pyramid vector quantisation (PVQ) [4] is form of lattice vector quantisation that generates fixed-length codes. PVQ has been used previously for still-image coding in conjunction with DCT [5] and wavelet [6] based techniques. In this paper, the performance of PVQ with both techniques is compared for a noisy channel. In [7], PVQ is used for contribution quality video coding at 22 Mbits/s. Here, PVQ is considered for low bit-rate video coding over noisy channels. In the next section, the theory behind PVQ is briefly reviewed. PVQ is then applied to still image coding using both DCT-based and wavelet-based techniques. The work is then extended to motion compensated video where PVQ is used for coding the prediction error. A novel fixed-rate video coder suitable for low bit-rate applications is presented and its resilience to random errors is compared to H.263.

2. PYRAMID VQ - THEORY

PVQ [4] is based on a source producing random variables having independent, identical Laplacian distributions. If these variables are grouped into vectors, X , of dimension L , it can be shown that for sufficiently large dimensions, vector X becomes highly localised around a particular hyperpyramid of dimension L . Since all points lying on the pyramid surface have equal probability, a pyramid vector quantiser can be constructed which consists of all points that lie both on a cubic lattice and on the hyperpyramid surface. However, for moderate sizes of vector dimension that are normally used in practice, significant distortion can be introduced by quantising on a single hyperpyramid. So, quantisation can be done on a number of concentric pyramids instead. This variation of PVQ is known as *product code* PVQ. The use of a number of concentric pyramids means that all codewords are no longer equiprobable. However, product code PVQ still gives good performance because of the possibility of using large vector dimensions compared to conventional VQ.

The PVQ codebook vectors will then consist of points on the surface of a number of concentric pyramids, with an equal number of points on each pyramid. The design bit-rate determines the number of points on each pyramid surface. Thus, a PVQ codeword consists of the pyramid radius r , which indicates on which concentric pyramid the vector lies, and the position of the vector on the pyramid. The pyramid radius, r , can be coded using an optimum scalar quantiser for the distribution.

In order to code the position of the vector on the pyramid, a unique fixed length binary index is assigned to every possible position using an indexing algorithm. The indexing algorithm presented in [6], and referred to as *product enumeration*, is used here. This index, together with the scalar quantised radius, is concatenated to form the PVQ codeword. This results in a fixed-length codeword for each codebook vector, making PVQ robust to channel errors.

3. ROBUST IMAGE CODING WITH PVQ

It follows directly from its derivation that PVQ is particularly suited for coding sources having a Laplacian distribution. It is known that the distribution of the coefficients resulting from transform and subband/wavelet coding of images can be modelled by a Laplacian distribution. In this section, DCT and wavelet based image coders using PVQ are described and their performance in error-free and random error conditions are compared.

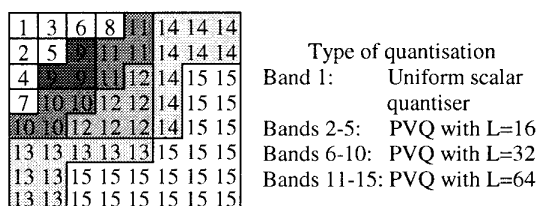


Figure 1: Partitioning of 8 by 8 block of DCT coefficients into 15 bands.

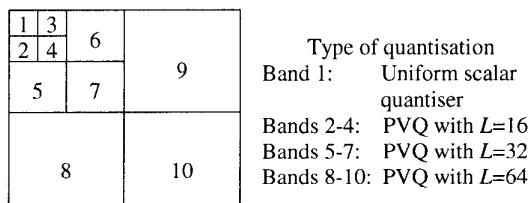


Figure 2: Wavelet decomposition of image into 10 subbands

3.1 DCT-based Coder

For the DCT-based coder, the original image is divided into non-overlapping blocks of 8 by 8 pixels. The DCT is then applied to each block, resulting in an 8 by 8 block of DCT coefficients. The coefficients are grouped into 15 different categories according to their location in the 8 by 8 block of DCT coefficients (figure 1). The coefficients in the same category in every 8 by 8 block over the entire image are then grouped together to form a *band*. Band numbering is as shown in Figure 1, e.g. band 1 consists of all dc coefficients. Each of the 15 bands is then coded with PVQ at a different rate and with a different vector dimension L , except for band 1, which is coded with a fixed-rate uniform scalar quantiser.

3.2 Wavelet-based Coder

For the wavelet-based coder, the (9,7) biorthogonal filters presented in [8] are used. A three-level hierarchical decomposition is applied, resulting in a total of 10 subbands, numbered as shown in Figure 2. The dc subband, i.e. band 1, is coded with a uniform fixed-rate scalar quantiser whereas the other subbands are quantised with PVQ at different rates and vector dimensions.

3.3 Bit-allocation and Quantisation

The dc band is essentially a low-pass version of the original image. For simplicity and robustness, it is quantised with a fixed rate uniform scalar quantiser. The bit rate allocated to the dc band is predefined and depends on the overall bit-rate. The bit-allocation algorithm for the remaining bands is based only on the weighted variance of the coefficients in the band and is derived from the optimum bit-allocation strategy for transform coding [9]. The bit-rate, R_k , for each band k is given by

$$R_k = R_T + \frac{1}{2} \log_2 \frac{\sigma_k^2 M_i}{\prod_{i=1}^N (\sigma_i^2 M_i)^{\frac{1}{M_i}}} \quad \text{for } k=2 \text{ to } N$$

where σ_k^2 = weighted variance of band k , R_T = total bit-rate.
 M_i = total number of coefficients / number of coefficients in band i .

The coefficients in each band are grouped into L -dimensional vectors, and these are then quantised with PVQ. PVQ assumes that the elements within the individual L -dimensional vectors are uncorrelated. Therefore the vectors are formed from coefficients spaced as far apart as possible within the band in order to ensure maximum decorrelation. The pyramid radius, r , is quantised with an optimum scalar quantiser for Laplacian distribution as used in [5]. Simulations with real images have shown that the radius distribution does not exactly match the Laplacian distribution, but very little is gained in terms of compression efficiency if a quantiser optimised to the radius distribution is used.

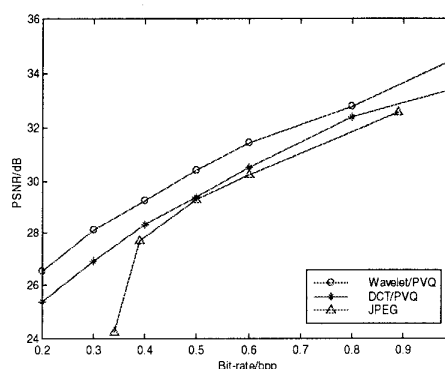


Figure 3: Compression performance of PVQ for Lenna.



Figure 4: (a) DCT/PVQ and (b) wavelet/PVQ at 0.5 bpp.

3.4 Compression Efficiency and Error-Resilience

The compression efficiency of our PVQ image coders for the 256 x 256 greyscale Lenna image is compared with JPEG in Figure 3, and decoded images at 0.5 bpp are shown in Figure 4. It is

assumed that the variances of each band are transmitted as header information and the required overhead for that is negligible. Our PVQ coder gives better PSNR than JPEG at all bit-rates, and more interestingly, the wavelet coder performs about 1 dB better than the DCT at all bit-rates.

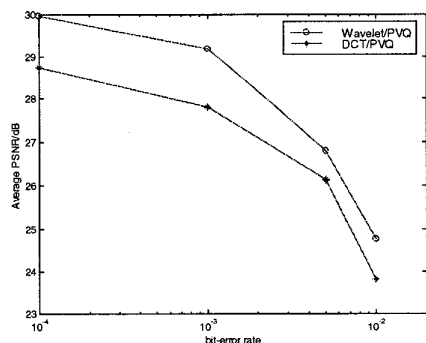


Figure 5: PVQ with random errors at 0.5 bpp.



Figure 6: (a) DCT/PVQ and (b) wavelet/PVQ at 0.5bpp and 1% BER.

Figure 5 shows simulation results in the presence of random bit-errors at various bit-error rates (BER) for Lenna at 0.5 bpp. A simple error concealment algorithm is used to detect errors in the dc band. For typical images, the dc coefficients are highly correlated with their spatial neighbours, i.e. any coefficient will have a similar value to at least one of its eight neighbours, unless it has been corrupted with errors. The error concealment algorithm looks at each dc coefficient and finds the minimum difference between that coefficient and its eight neighbours. If that difference exceeds a certain threshold, the coefficient is replaced by the average of its eight neighbours. Since a bit-error affects only a single codeword and does not propagate, PVQ provides a graceful degradation in image quality as the error rate increases. Figure 6 shows typical decoded images at 1% BER.

4. ROBUST VIDEO CODING WITH PVQ

In this section, a fixed-rate video coder based on the still-image PVQ/wavelet coder presented in section 3 is developed and its compression performance and robustness to errors is compared to H.263.

4.1 Wavelet/PVQ Video Coder

The structure of our PVQ based video coder is shown in Figure 7. It is based on H.263, except that motion compensation is done on overlapping blocks and the wavelet/PVQ still-image coder presented in the previous section is used to code the displaced frame difference (DFD).

Overlapped Block Motion Compensation: In H.263 without the advanced motion estimation mode, block-matching motion compensation is performed on non-overlapping 16 by 16 blocks with half-pixel accuracy. This often results in blocking artefacts in the DFD, which can adversely affect the performance of a wavelet coder. The use of overlapped motion compensation [10] can considerably improve the compression efficiency of a wavelet coder [11]. Therefore, overlapped motion compensation with a raised cosine window is used in our wavelet/PVQ video coder. Thus one motion vector for each 16 by 16 block is transmitted. In our simulations, full search block matching was used, but fast-search methods can be used with minimal loss in performance.

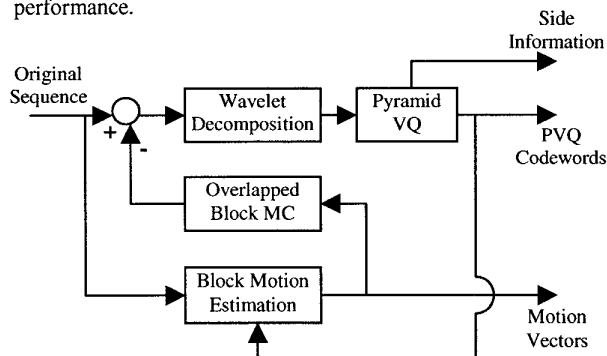


Figure 7: Block diagram of wavelet/PVQ video coder.

Coding of Motion Vectors: In H.263, the motion vector for a macroblock is predicted from the motion vectors of neighbouring macroblocks, and the prediction error is then variable length coded. This can cause a single bit-error in the motion vector to propagate from one macroblock to the next. In order to achieve maximum robustness, we code each motion vector separately with 12 bits, i.e. 6 bits for each component, giving a search area of $[-16, 15.5]$ pixels with half-pixel accuracy. This reduces the compression efficiency but gives greater robustness to errors and results in a video coder with a fixed bit-rate.

Overhead Information: The only overhead information required is the standard deviation of each band in order to perform the bit-allocation at the decoder. Simulation results for a number of sequences have shown that the 3 highest frequency bands is almost always allocated zero bits, and the standard deviation is not required for these. For the remaining bands, the standard deviation can be quantised with 7 bits each, giving a total of 35 bits/frame. This information is essential for the decoder to operate and must be protected with a powerful error-correcting code. Assuming a sequence at 5 frames/s and a (30,6) error-correcting code, this gives a total overhead of less than 1 kbps.

4.2 Compression Efficiency and Error-Resilience

The compression efficiency and robustness of our fixed-rate wavelet/PVQ coder has been compared with H.263, with all the options turned off. The same quantiser for all frames and a fixed frame rate of 5 frames/s is used. Results for 40 frames of the Foreman sequence (QCIF) are shown in Figure 8. Every 20th frame is intracoded for added robustness. For H.263, the frames are coded with a quantiser value giving an average bit-rate of 24 kbps. For the wavelet/PVQ coder, the intraframes are coded at

0.7 bpp whereas the DFDs are coded at 0.11 bpp, giving a total of about 24 kbps, including the overhead information. H.263 performs about 1dB better than our coder, mainly because of its more efficient coding of motion vectors.

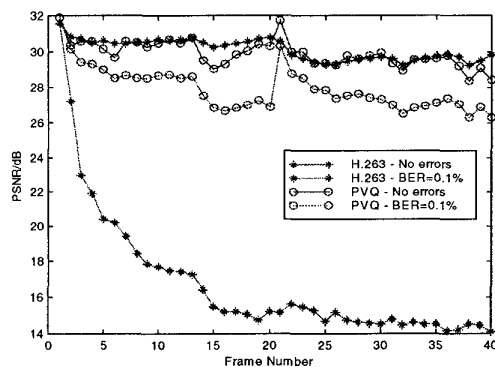


Figure 8: Wavelet/PVQ and H.263 at 24 kbps.

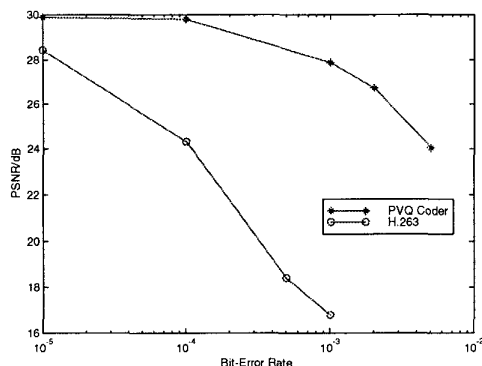


Figure 9: Wavelet/PVQ and H.263 with random errors.

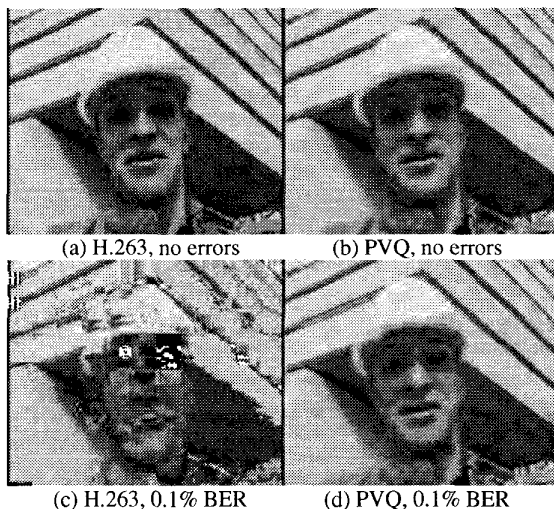


Figure 10: Frame 15 of Foreman with H.263 and Wavelet/PVQ at 24 kbps.

Simulations were carried out with random bit-errors in the compressed bitstream at various bit-error rates. For H.263, macroblocks that are undecodable because of errors are

concealed using the corresponding macroblock from the previous frame. The first frame, which is intra-coded, is assumed to be error-free. The results for one simulation with 0.1% BER is given in Figure 8. The PSNR decreases rapidly with H.263 as errors propagate, and even the intraframes are very sensitive to errors, whereas for our wavelet/PVQ coder, error propagation is minimised and the quality goes up when an intraframe is received. The mean PSNR for 40 frames of Foreman (QCIF, 5 Hz) for various bit-error rates and averaged over a number of simulations is shown in Figure 9. Typical decoded frames with 0.1% BER are shown in Figure 10 for subjective comparison.

5 CONCLUSIONS

In this paper, still image and video coders based on PVQ are developed that generate fixed length codewords only. Therefore they do not suffer from error propagation which is a severe problem when variable length coders such as H.263 are used in error prone environments. Since it generates a fixed bit-rate, the wavelet/PVQ video coder is particularly suitable for fixed rate channels. PVQ can be easily implemented as a low complexity algorithm with minimal power consumption [12]. This makes the wavelet/PVQ coder ideal for video transmission over mobile radio channels. The proposed wavelet/PVQ video coder performs slightly worse than H.263 at low-bit rates but is much more robust to random bit errors.

6. REFERENCES

- [1] Ferguson T.J. and Rabinowitz J.H., "Self-synchronizing Huffman codes," *IEEE Trans. Inf. Theory*, vol. IT-30, no. 4, July 1984.
- [2] Wen J.T. and Villasenor J.D., "A class of reversible variable length codes for robust image and video coding," *Proc. ICIP '97*, Santa Barbara, vol. 2, Oct. 1997.
- [3] Redmill D.W. and Kingsbury N.G., "The EREC: An error-resilient technique for coding variable-length blocks of data," *IEEE Trans. Image Proc.*, vol. 5, no. 4, April 1996.
- [4] Fischer T.R., "A pyramid vector quantizer," *IEEE Trans. Inform. Theory*, vol. IT-32, no. 4, July 1986, pp. 568-583.
- [5] Mohd-Yusof Z. and Fischer T.R., "An entropy-coded lattice vector quantizer for transform and subband image coding," *IEEE Trans. Image Proc.*, vol. 5, no. 2, Feb. 1996.
- [6] Hung A.C., Tsem E.K. and Meng T.H., "Error-resilient pyramid vector quantization for image compression," *IEEE Trans. Image Proc.*, vol. 7, no. 10, Oct. 1998.
- [7] D'Allessandro P. and Lancini R., "Video coding scheme using DCT-Pyramid Vector Quantization," *IEEE Trans. Image Proc.*, vol. 4, no. 3, March 1995.
- [8] Villasenor J.D. et al., "Wavelet filter evaluation for image compression," *IEEE Trans. Image Proc.*, vol. 4, no. 8, Aug. 1995.
- [9] Jayant N.S. and Noll P., *Digital Coding of Waveforms*, Prentice-Hall, 1984.
- [10] Auyeung C. et al., "Overlapped block motion compensation," *Proc. VCIP*, Boston, Nov. 1992, vol. 1818.
- [11] Mattavelli M., Nicoulin A. and Fernandez G., "Overlapped motion compensation for subband coding of video sequences," *Signal Processing: Image Communication*, vol. 8, 1996.
- [12] Meng T.H. et al., "Portable video-on-demand in wireless communication," *IEEE Proc.*, vol. 83, no. 4, April 1995.